

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : <b>C12P 21/00, 21/04, C12N 9/00 C12N 15/00, A01H 1/00</b>		A1	(11) International Publication Number: <b>WO 92/11382</b>									
			(43) International Publication Date: <b>9 July 1992 (09.07.92)</b>									
<p>(21) International Application Number: <b>PCT/US91/09654</b></p> <p>(22) International Filing Date: <b>19 December 1991 (19.12.91)</b></p> <p>(30) Priority data:</p> <table><tr><td>632,383</td><td>21 December 1990 (21.12.90)</td><td>US</td></tr><tr><td>731,226</td><td>16 July 1991 (16.07.91)</td><td>US</td></tr><tr><td>735,065</td><td>24 July 1991 (24.07.91)</td><td>US</td></tr></table> <p>(71) Applicant: CALGENE, INC. [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).</p> <p>(72) Inventors: SHEWMAKER, Christine, K. ; 1409 Spring Creek, Woodland, CA 95695 (US). STALKER, David, M. ; 2736 Cumberland Place, Davis, CA 95616 (US).</p> <p>(74) Agents: LASSEN, Elizabeth et al.; Calgene, Inc., 1920 Fifth Street, Davis, CA 95616 (US).</p>		632,383	21 December 1990 (21.12.90)	US	731,226	16 July 1991 (16.07.91)	US	735,065	24 July 1991 (24.07.91)	US	<p>(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), MC (European patent), NL (European patent), SE (European patent).</p> <p>Published <i>With international search report.</i></p>	
632,383	21 December 1990 (21.12.90)	US										
731,226	16 July 1991 (16.07.91)	US										
735,065	24 July 1991 (24.07.91)	US										
<p>(54) Title: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS</p> <p>(57) Abstract</p> <p>This invention relates to glycogen biosynthesis enzymes in plants. In particular, this invention is directed to plant cells having a DNA sequence encoding a glycogen biosynthesis enzyme integrated in its genome as the result of genetic engineering. Cells containing a DNA or RNA (mRNA) sequence encoding the enzyme as well as cells containing the enzyme are also provided. Plants and, more particularly, plant parts may also be obtained which contain glycogen biosynthesis enzyme sequences and/or containing such glycogen biosynthesis enzymes.</p>												

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU+	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TC	Togo
DE*	Germany	MC	Monaco	US	United States of America

+ Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.

GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS

This application is a continuation-in-part of USSN  
5 07/632,383 filed December 21, 1990 and a continuation-in-  
part of USSN unassigned filed on July 16, 1991.

Technical Field

The present invention is directed to compositions and  
10 methods related to sequences relevant to the biosynthesis  
of glycogen and their use in plants.

**INTRODUCTION**Background

15 In the animal kingdom, as well as in some lower  
plants, i.e., nonvascular plants, fungi, yeast and  
bacteria, the primary reserve polysaccharide is glycogen.  
Glycogen is a polysaccharide containing linear molecules  
with  $\alpha$ 1,4 glycosyl linkages and is branched via  $\alpha$ 1,6  
20 glycosyl linkages. Although from a linkage comparison,  
glycogen is analogous to starch, glycogen exhibits a  
different chain length and degree of polymerization. In  
bacteria, for example, the  $\alpha$ 1,6 glycosyl linkages  
constitute only approximately 10% of the total linkages,  
25 indicating that the majority of the glycogen polymer  
resides as a linear glucose molecule.

In plants, i.e. vascular plants, reserve  
polysaccharides are stored in roots, tubers and seeds in  
the form of starch. Starch, a complex polymer of glucose,

consists of a mixture of linear chain and branched chain glucans known as amylose and amylopectin respectively.

Starches isolated from different plants are found to have variable contents of amylose. Typically, amylose comprises

5 from about 10-25% of plant starch, the remainder being the branched polymer amylopectin. Amylopectin contains low molecular weight chains and high molecular weight chains, with the low molecular weight chains ranging from 5-30 glucose units and the high molecular weight chains from 30-  
10 100 or more. The ratio of amylose/amyopectin and the distribution of low molecular weight to high molecular weight chains in the amylopectin fraction are known to affect the properties, such as thermal stabilization, retrogradation, and viscosity, and therefore utility of  
15 starch. The highest published low m.w./high m.w. chain ratios (on a weight basis) in amylopectin are 3.9/1 for waxy corn starch which has unique properties.

Additionally, *duwx*, which has slightly more branch points than waxy has further unique properties.

20 In addition, starches from different plants or plant parts often have different properties. For example, potato starch has different properties than other starches, some of which may be due to the presence of phosphate groups. In some plant species, mutants have been identified which  
25 have altered contents of amylose and amylopectin.

Mutations that affect the activity of starch-branched enzyme in peas, for example, result in seeds having less starch and a lower proportion of amylopectin. Also, mutations in the *waxy* locus of maize, which encodes a

starch granule bound starch synthase, result in plants which produce amylopectin exclusively. Similarly, a potato mutant has been identified whose starch is amylose-free (Hovenkamp-Hermelink et al. *Theor. Appl. Genet.* (1987)

5 75:217-221). It has been found that varying the degree of starch branching can confer desirable physical properties; other changes in the characteristics of native starch could result in the production of polymers with new applications.

With the development of genetic engineering  
10 techniques, it is now possible to transfer genes from a variety of organism into the genome of a large number of different plant species. This process is preferable to plant breeding techniques whereby genes can only be transferred from one plant in a species to another plant in  
15 that same species or to a plant from a different, but closely related species. It would thus be desirable to develop plant varieties through genetic engineering, which have increased capacity for starch synthesis, altered amylose/amylopectin ratios, altered distribution of low to  
20 high molecular weight chains in the amylopectin fraction and also starches with novel molecular weight characteristics. In this manner, useful starches with a variety of viscosity or texture differences may be obtained.

25 To this end, nucleic acid sequences which encode glycogen biosynthetic enzymes are desirable for study and manipulation of the starch biosynthetic pathway. In particular, these enzymes may be expressed in plant cells

using plant genetic engineering techniques and targeted to a plastid where starch synthesis occurs.

Relevant Literature

5       The structural genes encoding the *E. coli* glycogen biosynthetic enzymes have been cloned (Okita, et al. (1981) *J. Biol. Chem.* 256:6944-6952) and their nucleic acid sequences determined (Preiss, J. (1984) *Ann. Rev. Microbiol.* 38:419-458; Kumar et al. (1986) *J. Biol. Chem.* 10 261:16256-16259). Genes encoding mammalian glycogen synthases have also been cloned and their nucleic acid sequences determined (Browner, et al. *Proc. Nat. Acad. Sci.* (1989) 86:1443-1447; Bai, et al., *J. Biol. Chem.* (1990) 265:7843-7848).

15

**BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 depicts a DNA sequence (SEQ ID NO: 1) for the *E. coli* glycogen synthase gene, *glgA*, generated through Polymerase Chain Reaction (PCR) from *E. coli* K-12 618.

20       Figure 2 depicts the translated amino acid sequence (SEQ ID NO:2) of the PCR generated *glgA* gene.

Figure 3 depicts DNA sequence (SEQ ID NO: 3) and the translated amino acid sequence of the PCR generated *E. coli* ADP-glucose pyrophosphorylase gene, *glgC*, from *E. coli* K-12 25 618.

**SUMMARY OF THE INVENTION**

By this invention, nucleic acid constructs comprising a glycogen biosynthetic enzyme sequence and at least one sequence endogenous to a plant are provided. As plants do not produce glycogen, these enzymes and sequences are not naturally found in a plant cell. Particularly, this invention relates to constructs comprising sequences relating to the glycogen biosynthetic enzymes, glycogen synthase and/or ADP-glucose pyrophosphorylase.

10 In one aspect of the invention, a sequence encoding a glycogen biosynthetic enzyme is joined to a sequence which encodes a transit peptide that provides for translocation of the glycogen biosynthetic enzyme to a plastid.

Other constructs of this invention provide sequences 15 for transcription of glycogen biosynthetic enzyme sequences in plant cells. To this end, transcriptional initiation regions that function to regulate expression of genes in plants are considered. Of particular interest are those regulatory regions that preferentially direct expression of 20 genes in roots, tubers, and seeds of plants or in other plant parts that synthesize reserve starch. In addition, constructs may contain sequences encoding a marker enzyme for selection of transformed cells.

Expression constructs which comprise sequences which 25 provide for transcriptional and translational regulation in plant cells of the sequences encoding glycogen biosynthetic enzymes are of special interest. These constructs include, in the 5' to 3' direction of transcription, a transcriptional/ translational initiation control region, a

sequence encoding a glycogen biosynthetic enzyme in reading frame, and a transcription/translation termination region, wherein the sequence encoding the enzyme is under the regulatory control of the initiation and termination regions. Expression constructs may also contain sequences which encode a transit peptide that provides for translocation of the glycogen biosynthesis enzymes to plastids and/or a marker enzyme.

In another aspect of the invention, vectors are considered which comprise sequences providing for transfer of desired sequences and integration into the genome of a plant cell. For example a plant transformation vectors may include *Agrobacterium* T-DNA border region(s) to provide for transfer of the sequences to the plant cell.

Also considered part of this invention are plant cells containing nucleic acid sequences of a glycogen biosynthetic enzyme. Such plant cells are obtainable through plant transformation techniques which utilize *Agrobacterium* to transfer DNA to the plant cells or alternatively through direct transfer techniques such as electroporation, microinjection or DNA bombardment. In either case, plant cells containing the desired sequences can be regenerated to yield whole plants containing the sequences.

In yet another aspect of this invention, plant cells containing the glycogen biosynthetic enzymes or having reduced or increased starch precursor enzymes are considered. Of particular interest are plant cells in starch storage organs, such as roots, tubers or seeds. It

is preferable that the enzyme be located in plastids, where starch synthesis occurs, and more preferably in amyloplasts, where reserve starch is synthesized and stored.

5       Further, it can be recognized that the modulation of glycogen biosynthetic enzymes in these plant cells has implications for modifying the starch content and/or composition of these cells. In this manner, plants or plant parts which synthesize and store starch may be  
10 obtained which have increased or decreased starch content and modified starch related properties such as specific gravity, free sugar content and/or novel and useful starches. In particular, potato starch having decreased amylose and modified amylopectin may be produced and  
15 further applications to modify starches consisting entirely of amylopectin such as that of waxy maize or a mutant potato, are also considered. Similarly, the starch from these plant parts can be harvested for use in commercial applications.

20

#### DETAILED DESCRIPTION OF THE INVENTION

This invention relates to glycogen biosynthesis enzymes in plants. In particular, this invention is directed to plant cells having a DNA sequence encoding a  
25 glycogen biosynthesis enzyme integrated in its genome as the result of genetic engineering. Cells containing a DNA or RNA (mRNA) sequence encoding the enzyme as well as cells containing the enzyme are also provided. Plants and, more particularly, plant parts may also be obtained which

contain glycogen biosynthesis enzyme sequences and/or containing such glycogen biosynthesis enzymes.

The biosynthetic steps involved in glycogen synthesis in *E. coli* include: 1) the formation of ADP-glucose from 5 ATP and glucose 1-phosphate, 2) the transfer of a glucose unit from ADP-glucose to a preformed maltodextrin primer via an  $\alpha$ 1,4 linkage, and 3) the formation of  $\alpha$ 1,6 glucosyl linkages from glycogen. The bacterial enzymes which catalyze the above reactions are ADP-glucose 10 pyrophosphorylase (EC 2.7.7.27), glycogen synthase (EC 2.4.1.21), and Q-enzyme or branching enzyme (EC 2.4.1.18), respectively. The genes encoding these enzymes have been cloned and are also known as *glgC*, *glgA*, and *glgB*, respectively.

15 The pathway of glycogen biosynthesis in mammals is similar to that seen in bacteria, an exception being that UDP-glucose is the preferred glucose donor. The mammalian enzymes which catalyze glycogen biosynthetic reactions similar to those in bacteria are glucose-1-phosphate 20 uridylyltransferase, glycogen synthase (EC 2.4.1.11), and 1,4- $\alpha$ -glucan branching enzyme. Genes encoding a human muscle and a rat liver glycogen synthase have been cloned and their nucleic acid sequences determined.

In particular, the glycogen biosynthesis enzyme 25 glycogen synthase (*glgA*) is of special interest. The *E. coli* glycogen synthase is of particular interest in that the enzyme is similar to plant starch synthase with respect to being non-responsive to allosteric effectors or chemical modifications. Expression of a glycogen synthase enzyme in

a plant host demonstrates biological activity even within an intact plant cell. Namely, potato plants having *glgA* expressed in potato tubers result in tubers having a decreased specific gravity; specific gravity being a 5 commonly used measurement with respect to dry matter and starch contents of potato tubers (W.G. Burton, in *The Potato*, Third Edition, pub. Longman Scientific & Technical (1989) Appendix II, pp. 599-601). Further analysis of transgenic tubers having decreased specific gravity 10 indicates that the starch in these tubers is modified. In particular, the percentage of amylose is decreased and the ratio of low m.w./high m.w. chains in the amylopectin fraction is increased. This phenotypic effect *in planta* is indicative of *glgA* biological activity.

15 Other phenotypic starch modifications resulting from biological activity of glycogen biosynthetic enzymes in plants are also considered in this invention. Such altered phenotypes may result from enzymatic activity of these proteins on plant starch precursors, or alternatively from 20 the inhibition of plant starch biosynthetic enzyme activities. Inhibition of plant enzymes, for example, could result through the production of inactive forms of the plant enzymes as the result of association with the glycogen biosynthetic enzymes. The inhibition of plant 25 enzymes may then lead to plants having altered starch (such as branching patterns or molecular weight) and/or lowered starch levels. In addition, increased plant metabolites, such as sugars, could also result from starch alteration or inhibition caused by expression of glycogen biosynthetic

enzymes. For example, transgenic potato tubers described herein are observed to have up to 3-fold increases in free sugar content.

Measurement of specific gravity or free sugar content 5 may be useful to detect modified starch, with other methods, such as HPLC and gel filtration, also being useful. The glycogen synthase sequence may be employed as the sole glycogen biosynthetic enzyme or in conjunction with sequences encoding other glycogen biosynthetic 10 enzymes.

An enzyme relevant to the biosynthesis of glycogen is considered in this invention as including any sequence of amino acids, such as protein, polypeptide, or peptide fragment, which demonstrates the ability to catalyze a 15 reaction involved in the biosynthesis of glycogen. Thus, a glycogen biosynthetic enzyme of this invention will display activity towards a glucan molecule, although it may have preferential activity towards either ADP- or UDP-glucose. In plants, ADP-glucose is the preferred donor for starch 20 biosynthetic reactions. Therefore, of particular interest in this invention are glycogen biosynthesis enzymes which also prefer ADP-glucose. Of special interest are glycogen biosynthesis enzymes obtainable from bacterial sources. Over 40 species of bacteria synthesize glycogen, including 25 *Escherichia* and *Salmonella*.

Obtaining glycogen biosynthetic enzymes may be accomplished by a variety of methods known to those skilled in the art. For example, radiolabeled nucleic acid probes may be prepared from a known sequence which will bind to

and thus provide for detection of other sequences, glycogen biosynthesis proteins may be purified and sequences obtained through biochemical or antibody techniques, polymerase chain reaction (PCR) may be employed based upon known nucleic acid sequences, and the like.

Typically, a sequence obtainable from the use of nucleic acid probes will show 60-70% sequence identity between the target sequence and the given sequence encoding a glycogen biosynthetic enzyme of interest. However,

lengthy sequences with as little as 50-60% sequence identity may also be obtained. The nucleic acid probes may be a lengthy fragment of the nucleic acid sequence, or may also be a shorter, oligonucleotide probe. Oligonucleotide probes can be considerably shorter than the entire nucleic acid sequence encoding a glycogen biosynthetic enzyme, but should be at least about 10, preferably at least about 15, and more preferably at least about 20 nucleotides. A higher degree of sequence identity is desired when shorter regions are used as opposed to longer regions. It may thus be desirable to identify enzyme active sites where amino acid sequence identity is high to design oligonucleotide probes for detecting homologous genes.

When longer nucleic acid fragments are employed as probes (greater than about 100 bp), one may screen at lower stringencies in order to obtain sequences from the target sample which have 20-50% deviation (i.e., 50-80 sequence homology) from the sequences used as probe. Hybridization and washing conditions can be varied to optimize the hybridization of the probe to the sequences of interest.

Lower temperatures and higher salt (SSC) concentrations allow for hybridization of more distantly related sequences (low stringency). If background hybridization is a problem under low stringency conditions, the temperature can be

5 raised either in the hybridization or washing steps and/or salt content lowered to improve detection of the specific hybridizing sequence. Hybridization and washing temperatures can be adjusted based on the estimated melting temperature of the probe. (See, for example, Beltz, et al.

10 *Methods in Enzymology* (1983) 100:266-285).

It will be recognized by one of ordinary skill in the art that glycogen biosynthetic enzyme sequences of this invention may be modified using standard techniques of site specific mutation or PCR, or modification of the sequence

15 may be accomplished in producing a synthetic nucleic acid sequence and will still be considered a glycogen biosynthesis enzyme nucleic acid sequence of this invention. For example, wobble positions in codons may be changed such that the nucleic acid sequence encodes the

20 same amino acid sequence, or alternatively, codons can be altered such that conservative amino acid substitutions result. In either case, the peptide or protein maintains the desired enzymatic activity and is thus considered part of the instant invention.

25 A nucleic acid sequence of an enzyme relevant to the biosynthesis of glycogen of this invention may be a DNA or RNA sequence, derived from genomic DNA, cDNA, mRNA, or may be synthesized in whole or in part. The structural gene sequences may be cloned, for example, by isolating genomic

DNA from an appropriate source, and amplifying and cloning the sequence of interest using a polymerase chain reaction (PCR). Alternatively, the gene sequences may be synthesized, either completely or in part, especially where 5 it is desirable to provide plant-preferred sequences.

Thus, all or a portion of the desired structural gene may be synthesized using codons preferred by a selected plant host. Plant-preferred codons may be determined, for example, from the codons used most frequently in the 10 proteins expressed in a particular plant host species.

Other modifications of the gene sequences may result in mutants having slightly altered activity. Once obtained, a glycogen biosynthetic enzyme nucleic acid sequence of this invention may be combined with other sequences in a variety 15 of ways.

Often, the sequences associated with glycogen biosynthesis are used in conjunction with endogenous plant sequences. By "endogenous plant sequence" is meant any sequence which can be naturally found in a plant cell. 20 These sequences include native (indigenous) plant sequences as well as sequences from plant viruses or plant pathogenic bacteria, such as *Agrobacterium* or *Rhizobium* species that are naturally found and functional in plant cells.

In one aspect of this invention, the glycogen biosynthesis enzyme will be joined to a sequence encoding a transit peptide or functional portion of a transit peptide which is capable of providing for intracellular transport 25 of a heterologous protein to a plastid in a plant host cell. Chloroplasts are the primary plastid in

photosynthetic tissues, although plant cells are likely to have other kinds of plastids, including amyloplasts, chromoplasts, and leucoplasts. Transport into amyloplasts is preferred in this invention as these plastids are 5 associated with reserve starch synthesis and storage. Any transit peptide providing for intracellular transport to a plastid is useful in this invention, such as the transit peptides from the precursor proteins of the small subunit of ribulose bisphosphate carboxylase (RUBISCO), acyl 10 carrier protein (ACP), the waxy locus of maize, or other nuclear-encoded plastid proteins.

In addition to the identified transit peptide portion of a protein, it may be desirable to include sequences encoding a portion of the mature plastid-targeted protein 15 to facilitate the desired intracellular transport of the glycogen biosynthetic enzyme. In one embodiment of this invention, the transit peptide from the small subunit of RUBISCO is utilized along with 48 bp of sequence encoding the amino terminal 16 amino acids of a mature small subunit 20 protein.

Other endogenous plant sequences may be provided in nucleic acid constructs of this invention, for example to provide for transcription of the glycogen biosynthetic enzyme sequences. Transcriptional regulatory regions are 25 located immediately 5' to the DNA sequences of the gene of interest, and may be obtained from sequences available in the literature, or identified and characterized by isolating genes having a desirable transcription pattern in plants, and studying the 5' nucleic acid sequences.

Numerous transcription initiation regions which provide for a variety of constitutive or regulatable, e.g. inducible, expression in a plant cell are known. Among sequences known to be useful in providing for constitutive gene expression are regulatory regions associated with *Agrobacterium* genes, such as for nopaline synthase (*Nos*), mannopine synthase (*Mas*), or octopine synthase (*Ocs*), as well as regions coding for expression of viral genes, such as the 35S and 19S regions of cauliflower mosaic virus (CaMV). The term constitutive as used herein does not necessarily indicate that a gene is expressed at the same level in all cell types, but that the gene is expressed in a wide range of cell types, although some variation in abundance is often detectable. Other useful transcriptional initiation regions preferentially provide for transcription in certain tissues or under certain growth conditions, such as those from napin, seed or leaf ACP, the small subunit of RUBISCO, patatin, zein, and the like.

Sequences to be transcribed are located 3' to the plant transcription initiation region and may be oriented, in the 5'-3' direction, in the sense orientation or the antisense orientation. In the sense orientation, an mRNA strand is produced which encodes the desired glycogen biosynthetic enzyme, while in antisense constructs, an RNA sequence complementary to an enzyme coding sequence is produced. The sense orientation is desirable when one wishes to produce the glycogen biosynthetic enzyme in plant cells, whereas the antisense strand may be useful to

inhibit production of a related plant starch biosynthesis enzymes. Regions of homology have been observed, for example, upon comparison of *E. coli* *glgC* sequence to that of a rice ADP glucose pyrophosphorylase. Either method may 5 be useful in obtaining an alteration in the starch or dry matter content of a plant. The presence of glycogen biosynthetic enzyme sequences in the genome of the plant host cell may be confirmed, for example by a Southern analysis of DNA or a Northern analysis of RNA sequences or 10 by PCR methods.

In addition to sequences providing for transcriptional initiation in a plant cell, also of interest are sequences which provide for transcriptional and translational initiation of a desired sequence encoding a glycogen 15 biosynthetic enzyme. Translational initiation regions may be provided from the source of the transcriptional initiation region or from the gene of interest. In this manner, expression of the glycogen biosynthetic enzyme in a plant cell is provided. The presence of the glycogen 20 biosynthetic enzyme in the plant host cell may be confirmed by a variety of methods including a immunological analysis of the protein (e.g. Western or ELIZA), as a result of phenotypic changes observed in the cell, such as altered starch content, altered starch branching, etc., or by assay 25 for increased enzyme activity, and the like. If desired the enzyme may be harvested from the plant host cell or used to study the effect of the enzyme on plant cell functions, especially in the plastid organelles.

Other sequences may be included in the nucleic acid construct providing for expression of the glycogen biosynthetic enzymes ("expression constructs") of this invention, including endogenous plant transcription 5 termination regions which will be located 3' to the desired glycogen biosynthetic enzyme encoding sequence. In one embodiment of this invention, transcription termination sequences derived from a patatin gene are preferred. Transcription termination regions may also be derived from 10 genes other than those used to regulate the transcription in the nucleic acid constructs of this invention. Transcription termination regions may be derived from a variety of different gene sequences, including the *Agrobacterium*, viral and plant genes discussed above for 15 their desirable 5' regulatory sequences.

Further constructs are considered which provide for transcription and/or expression of more than one glycogen biosynthetic enzyme. For example, one may wish to provide enzymes to plant cells which provide for modification of 20 the starch synthesized, as well as for an increase or decrease in overall starch production. Examples of enzymes which may prove useful in modifying starch structure are those which catalyze reactions involving UDP- or ADP-glucose, for example glycogen synthase or branching enzyme. 25 However, to provide for increased or decreased starch production, one may wish to utilize sequences encoding enzymes which catalyze formation of the nucleotide-glucose molecule, such as ADP-glucose pyrophosphorylase in bacteria, or glucose-1-phosphate uridylyltransferase in

mammals. Although plants typically utilize ADP-glucose, UDP-glucose may also be useful.

In providing for transcription and/or expression of the glycogen biosynthetic enzyme sequences, one may wish to 5 limit these enzymes to plant cells which synthesize and store reserve starch. Towards this end, one can identify useful transcriptional initiation regions that provide for expression preferentially in the roots, tubers, seeds, or other starch-containing tissues of a desired plant species. 10 These sequences may be identified from cDNA libraries using differential screening techniques, for example, or may be derived from sequences known in the literature. Of particular interest in a presently preferred embodiment of the invention is a transcriptional initiation region from 15 the patatin gene of potato, which demonstrates preferential expression in the potato tuber. Similarly, other promoters which are preferentially expressed in the starch-containing tissues, such as the zein genes in corn, as opposed to other plant structures are desirable.

20 In developing the nucleic acid constructs of this invention, the various components of the construct or fragments thereof will normally be inserted into a convenient cloning vector, e.g. a plasmid, which is capable of replication in a bacterial host, e.g. *E. coli*. Numerous 25 vectors exist that have been described in the literature, many of which are commercially available. After each cloning, the cloning vector with the desired insert may be isolated and subjected to further manipulation, such as restriction, insertion of new fragments or nucleotides,

ligation, deletion, mutation, resection, etc. so as to tailor the components of the desired sequence. Once the construct has been completed, it may then be transferred to an appropriate vector for further manipulation in 5 accordance with the manner of transformation of the host cell.

The constructs of this invention providing for transcription and/or expression of glycogen biosynthetic enzyme sequences of this invention may be utilized as 10 vectors for plant cell transformation. The manner in which nucleic acid sequences are introduced into the plant host cell is not critical to this invention. Direct DNA transfer techniques, such as electroporation, microinjection or DNA bombardment may be useful. To aid in 15 identification of transformed plant cells, the constructs of this invention may be further manipulated to include plant selectable markers. The use of plant selectable markers is preferred in this invention as the amount of experimentation required to detect plant cells is greatly 20 reduced when a selectable marker is expressed. Useful selectable markers include enzymes which provide for resistance to an antibiotic such as gentamycin, hygromycin, kanamycin, and the like. Similarly, enzymes providing for production of a compound identifiable by color change, such 25 as GUS, or luminescence, such as luciferase are useful.

An alternative method of plant cell transformation employs plant vectors which contain additional sequences which provide for transfer of the desired glycogen biosynthetic enzyme sequences to a plant host cell and

stable integration of these sequences into the genome of the desired plant host. Selectable markers may also be useful in these nucleic acid constructs to provide for differentiation of plant cells containing the desired 5 sequences from those which have only the native genetic material. Sequences useful in providing for transfer of nucleic acid sequences to host plant cells may be derived from plant pathogenic bacteria, such as *Agrobacterium* or *Rhizogenes*, plant pathogenic viruses, or plant transposable 10 elements.

When *Agrobacterium* is utilized for plant transformation, it may be desirable to have the desired nucleic acid sequences bordered on one or both ends by T-DNA, in particular the left and right border regions, and 15 more particularly, at least the right border region. These border regions may also be useful when other methods of transformation are employed.

Where *Agrobacterium* or *Rhizogenes* sequences are utilized for plant transformation, a vector may be used 20 which may be introduced into an *Agrobacterium* host for homologous recombination with the T-DNA on the Ti- or Ri- plasmid present in the host. The Ti- or Ri- containing the T-DNA for recombination may be armed (capable of causing gall formation), or disarmed (incapable of causing gall 25 formation), the latter being permissible so long as a functional complement of the vir genes, which encode trans-acting factors necessary for transfer of DNA to plant host cells, is present in the transformed *Agrobacterium* host. Using an armed *Agrobacterium* strain can result in a mixture

of normal plant cells, some of which contain the desired nucleic acid sequences, and plant cells capable of gall formation due to the presence of tumor formation genes. Cells containing the desired nucleic acid sequences, but lacking tumor genes can be selected from the mixture such that normal transgenic plants may be obtained.

In a preferred method where *Agrobacterium* is used as the vehicle for transforming host plant cells, the expression or transcription construct bordered by the T-DNA border region(s) will be inserted into a broad host range vector capable of replication in *E. coli* and *Agrobacterium*, there being broad host range vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, et al., (*Proc. Nat. Acad. Sci., U.S.A.* (1980) 77:7347-7351) and EPA 0 120 515, which are incorporated herein by reference. Alternatively, one may insert the sequences to be expressed in plant cells into a vector containing separate replication sequences, one of which stabilizes the vector in *E. coli*, and the other in *Agrobacterium*. See, for example, McBride and Summerfelt (*Plant Mol. Biol.* (1990) 14:269-276), wherein the pRiHRI (Jouanin, et al., *Mol. Gen. Genet.* (1985) 201:370-374) origin of replication is utilized and provides for added stability of the plant expression vectors in host *Agrobacterium* cells.

Utilizing vectors such as those described above, which can replicate in *Agrobacterium* is preferred. In this manner, recombination of plasmids is not required and the host *Agrobacterium vir* regions can supply trans-acting

factors required for transfer of the T-DNA bordered sequences to plant host cells.

In general, the plant vectors of this invention will contain glycogen biosynthetic enzyme sequence(s), alone or 5 in combination with transit peptides, and endogenous plant sequences providing for transcription or expression of these sequences in a plant host cell. The plant vectors containing the desired sequences may be employed with a wide variety of plant cells, particularly plants which 10 produce and store reserve starch. Plants of interest include, but are not limited to corn, cereal grains, sorghum, rice, potato, tapioca, cassava, arrowroot and sago.

Also considered part of this invention are plants 15 containing the nucleic acid sequences of this invention, and following from that, plants containing glycogen biosynthetic enzymes as the result of expression of the sequences of this invention in plant cells or having a decreased expression of a native starch biosynthesis 20 enzyme. Methods of regenerating whole plants from plant cells are known in the art, and the method of obtaining transformed and regenerated plants is not critical to this invention. In general, transformed plant cells are cultured in an appropriate medium, which may contain 25 selective agents such as antibiotics, where selectable markers are used to facilitate identification of transformed plant cells. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the

shoots transferred to rooting medium for regeneration of plants. The plants may then be used to establish repetitive generations either from seed or using vegetative propagation techniques.

5 Of particular interest are plant parts, e.g. tissues or organs, (and corresponding cells) which form and store reserve starch, such as roots, tubers, and seeds. Of more particular interest are potato tubers containing the glycogen biosynthetic enzymes. It can be recognized that  
10 the modification of glycogen biosynthetic enzymes in plants may also result in desirable alterations in the plant cells or parts. These alterations may include modification of dry matter content, free sugar content or of starch content and/or structure, or modification of specific gravity. The  
15 novel plant cells or plant parts can thus be harvested and used for isolation of the altered material.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration  
20 only and are not intended to limit the present invention.

#### EXAMPLES

Example 1 Cloning of Glycogen Biosynthetic Enzyme  
25 Genes

A. Cloning and Sequencing of a GlgA Gene From E.  
coli

Total genomic DNA is prepared from *E. coli* K12 618  
(Leung et al., *J. of Bacteriology* (1986) 167:82-88) by  
30 growing a 5ml culture in ECLB (Maniatis, T. et al.,

*Molecular Cloning: A Laboratory Manual*, (1982) Cold Spring Harbor, NY) overnight at 37°C. The bacteria are pelleted by centrifugation for 10 minutes at 4500xg, the supernatant is discarded, and the pellet is resuspended in 2.5ml of 5 10mM Tris, 1mM EDTA buffer. To this suspension is added 500µl of a 5mg/ml Pronase® protease (Calbiochem Brand Biochemicals; La Jolla, CA) solution and 2ml of 2% lauryl sulfate, sodium salt (Sigma; St. Louis, MO), with gentle mixing, and the suspension is incubated at 37°C for 50 10 minutes. A clear solution indicates that the bacteria have lysed. The solution is then extracted with 5ml phenol, then 5ml phenol:chloroform:isoamyl alcohol (25:24:1), followed by 5ml chloroform. Nucleic acids are precipitated from the aqueous phase with 1/10 volume of 3M sodium 15 acetate and two volumes of 100% ethanol, and the tube is incubated at room temperature for 1 hour. Nucleic acids are removed from solution and resuspended in 1ml water. A second ethanol precipitation is performed and the nucleic acids are resuspended in 200µl of 10mM Tris, 1mM EDTA 20 buffer.

Synthetic oligonucleotides, str1 and str2, corresponding to sequences flanking the 1.4 kb *glgA* (glycogen synthase - EC 2.4.1.21) gene of *E. coli* (Kumar et al., *J. of Biol. Chemistry* (1986) 261:16256-16259) and 25 containing restriction sites for *Bgl*II (str1) and *Sal*I (str2) are synthesized on an Applied Biosystems 380A DNA synthesizer (Foster City, CA) in accordance with the manufacturer's instructions.

The nucleic acid preparation of *E. coli* is treated with RNase and the DNA is used as a template in a polymerase chain reaction (PCR) with str1 and str2 as primers. A Perkin-Elmer/Cetus (Norwalk, CT) thermal cycler 5 is used with the manufacturer's reagents and in accordance with the manufacturer's instructions. The reaction mixture contains 41.5 $\mu$ l H<sub>2</sub>O, 10 $\mu$ l 10X reaction buffer, 16 $\mu$ l dNTP's [1.25mM dCTP, dATP, dGTP & dTTP], 5 $\mu$ l str1 (20mM), 5 $\mu$ l str2 (20mM), 22 $\mu$ l total *E. coli* DNA (0.05 $\mu$ g/ $\mu$ l), and 0.5 $\mu$ l Tag 10 polymerase. The reaction is performed for 15 cycles with melting (denaturation) for 1 minute at 94°C, annealing (hybridization) for 2 minutes at 37°C and chain elongation for 3 minutes at 72°C. The reaction is then performed for an additional 10 cycles with melting for 1 minute at 94°C, 15 annealing for 2 minutes at 37°C and chain elongation at 72°C for 3 minutes 15 seconds initially and increasing the time by 15 seconds each cycle so that the last cycle is 5 minutes 45 seconds.

The resulting PCR products (~1.4 kb) are digested with 20 *Bgl*II and *Sal*I and ligated into a *Sal*I and *Bgl*II digest of pCGN789, a pUC based vector similar to pUC119 with the normal polylinker replaced by a synthetic linker which contains the restriction digest sites *Eco*RI, *Sal*I, *Bgl*II, *Pst*I, *Xho*I, *Bam*HI, and *Hind*III. The ligated DNA is 25 transformed into *E. coli* DH5 $\alpha$ . The transformed cells are plated on ECLB containing penicillin (300mg/L), IPTG and X-Gal (Vieira and Messing Gene (1982) 19:259-268). White colonies are picked to ECLB containing penicillin (300mg/L) and flooded with I<sub>2</sub>/KI (0.2% I<sub>2</sub> in 0.4% KI). Clones

producing a brown color, which indicates excess starch production, are selected. One clone, *glgA-2*, was selected and the DNA and translated amino acid sequences determined (see, Figs. 1 and 2 and SEQ ID NOS: 1-2). The DNA sequence 5 is 98% homologous to the published sequence (Kumar et al, *supra*) and 96% homologous at the amino acid level.

B. Cloning and Sequencing of a GlgC Gene From E. coli

Synthetic oligonucleotides, *glgC1* and *glgC2*, 10 corresponding to sequences flanking the 1.3 kb *glgC* (ADP-glucose pyrophosphorylase - EC 2.7.7.27) gene of *E. coli* (Baecker et al., *J. of Biol. Chemistry* (1983) 258:5084-5088) and containing restriction sites for *BglII* (*glgC1*) and *SalI* (*glgC2*) are synthesized on an Applied Biosystems 15 380A DNA synthesizer (Foster City, CA) in accordance with the manufacturer's instructions.

Total genomic DNA is prepared from *E. coli* K12 618 as described above. The nucleic acid preparation of *E. coli* is treated with RNase and the DNA is used as a template in 20 a polymerase chain reaction (PCR) with *glgC1* and *glgC2* as primers. A Perkin-Elmer/Cetus (Norwalk, CT) thermal cycler is used with the manufacturer's reagents as described above.

The resulting PCR products (~1.3 kb) are digested with 25 *BglII* and *SalI* and ligated into a *SalI* and *BglII* digest of pCGN789 (described above). The ligated DNA is transformed into *E. coli* DH5 $\alpha$ , and the transformed cells are plated as described above. Clones producing excess starch are selected as described above. One clone, pGlc-37, was

selected and the DNA sequence (SEQ ID NO: 3) determined (see, Fig. 3). The DNA sequence is 99% homologous to the published sequence (Baecker et al, *supra*) of *glgC* from *E. coli* K-12. The *glgC* from *E. coli* 618 is a mutant and the 5 amino acid sequence of this mutant differs from that of *E. coli* K-12 at five amino acids (Lee et al., *Nucl. Acids Res.* (1987) 15:10603). The translated amino acid sequence of pGlgC-37 differs from that of the *glgC* from *E. coli* 618 at a single amino acid; the asparagine (Asn) at position 361 10 of the *E. coli* 618 mutant is an aspartate (Asp) in the translated amino acid sequence of pGlgC-37 (Fig. 3).

**Example 2      Attachment of Glycogen Genes to SSU Leader Sequence**

15        A.    Construction of SSU + aroA Transit Peptide  
Plasmid pCGN1132 contains a 35S promoter, ribulosebisphosphate carboxylase small subunit (5'-35S-SSU) leader from soybean plus 48 bp of mature small subunit (SSU) gene from pea, and *aroA* sequence (the gene locus 20 which encodes 5-enolpyruvyl-3-phosphoshikimate synthetase (EC 2.5.1.19)). It is prepared from pCGN1096, a plasmid containing a hybrid SSU gene, which carries DNA encoding mature SSU protein from pea, and *SstI* and *EcoRI* sites 3' of the coding region (used in the preparation of pCGN1115, a 25 plasmid having a 5'-35S-SSU+48-*aroA*-tm1-3' sequence) and pCGN1129, (a plasmid having a 35S promoter in a chloramphenicol resistance gene (*Cam*<sup>R</sup>) backbone).

Construction of pCGN1096

The *aroA* moiety of pCGN1077 is removed by digestion 30 with *SphI* and *SalI*. (The construction of pCGN1077 and

other constructs hereunder are described in detail in co-pending U.S. application Serial No. 06/097,498, filed September 16, 1987, which is hereby incorporated by reference). In its place is cloned the region coding for 5 the mature pea SSU protein, as an *SphI-PstI* fragment, which is then excised with *SphI* and *SalI*. The resulting plasmid, pCGN1094, codes for a hybrid SSU protein having the transit peptide of the soybean clone, and the mature portion of the pea clone and contains *SstI* and *EcoRI* sites 3' of the 10 coding region. The *HindIII* to *BamHI* region of transposon Tn6 (Jorgensen et al., *Mol. Gen. Genet.* (1979) 177:65) encoding the kanamycin resistance gene (*Kan<sup>r</sup>*) is cloned into the same sites of pBR322 (Bolivar et al., *Gene* (1977) 2:95-133) generating pDS7. The *BglII* site 3' of the *Kan<sup>r</sup>* 15 gene is digested and filled in with the large fragment of *E. coli* DNA polymerase I and deoxynucleotides triphosphate. An *SstI* linker is ligated into the blunted site, generating plasmid pCGN1093. Plasmid pPMG34.3 is digested with *SalI*, the site filled in as above and *EcoRI* linkers are ligated 20 into the site resulting in plasmid pCGN1092. The latter plasmid is digested with *SstI* and *SmaI* and the *Kan<sup>r</sup>* gene excised from pCGN1093 with *SstI* and *SmaI* is ligated in, generating pCGN1095. The *Kan<sup>r</sup>* and *aroA* genes are excised 25 as a piece from pCGN1095 by digestion with *SstI* and *EcoRI* and inserted into the *SstI* and *EcoRI* sites of pCGN1094, producing pCGN1096. Summarizing, pCGN1096 contains (5' - 3') the following pertinent features: The SSU gene - a polylinker coding for *PstI*, *SalI*, *SstI*, and *KpnI* - the *Kan<sup>r</sup>*

gene - *Sma*I and *Bam*HI restriction sites - the *aroA* gene without the original ATG start codon.

Construction of pCGN1115

Plasmid pCGN1096 is digested to completion with *Sal*I and then digested with exonuclease *Bal*31 (BRL; Gaithersburg, MD) for 10 minutes, thus deleting a portion of the mature SSU gene. The resulting plasmid is then digested with *Sma*I to eliminate the *Kan*<sup>r</sup> gene and provide blunt ends, recircularized with T4 DNA ligase and transformed into *E. coli* LC3 (Comai et al., *Science* (1983) 221:370-371), an *aroA* mutant. DNA isolated from *aroA*<sup>+</sup> and *Kan*<sup>r</sup> colonies is digested with *Bam*HI and *Sph*I and ligated with *Bam*HI- and *Sph*I-digested M13mp18 (Norrrander et al., *Gene* (1983) 26:101-106 and Yanisch-Perron et al., *Gene* (1985) 33:103-119) DNA for sequencing. Clone 7 has 48 bp of the mature SSU gene remaining and the 3' end consists of phe-glu-thr-leu-ser (SEQ ID NO: 4). Clone 7 is transformed into *E. coli* strain 71-18 (Yanish-Perron et al. (1985) *supra*) and DNA isolated from transformants is digested with *Sph*I and *Cla*I to remove the 0.65 kb fragment containing the 48 bp of mature protein and the 5' end of the *aroA* gene. Plasmid pCGN1106 (Comai, L. et al., *J. Biol. Chem.* (1988) 263:15104-15109) is also digested with *Sph*I and *Cla*I and the 6.8 kb isolated vector fragment is ligated with the 0.65 kb fragment of clone 7 to yield pCGN1115 (5'-35S-SSU+48-*aroA*-tml-3').

Construction of pCGN1129

The 7.2 kb plasmid pCGN1180 (35S-SSU+70-*aroA*-ocs3') (Comai et al. (1988) *supra*) and the 25.6 kp plasmid pCGN594

(LB-Gent<sup>r</sup>-ocs5'-Kan<sup>r</sup>-ocs3'-RB) (construction of pCGN594 is described in co-pending U.S. application Serial No. 07/382,802, filed July 19, 1989) are digested with *Hind*III and ligated together to yield the 32.8 kb plasmid pCGN1109  
5 (LB-Gent<sup>r</sup>-35S-SSU+70-aroA-ocs3'-ocs5'-Kan<sup>r</sup>-ocs3'-RB).

Plasmid pCGN1109 is digested with *Eco*RI to delete an internal 9.1 kb fragment containing the SSU leader plus 70 bp of the mature SSU gene, the aroA gene and its ocs3' terminator, the Amp<sup>r</sup> backbone from pCGN1180 and ocs5'-Kan<sup>r</sup>-10 ocs3' from pCGN594. The *Eco*RI digest of pCGN1109 is then treated with Klenow fragment to blunt the ends, and an *Xba*I linker (dCCTCGAGG) (New England Biolabs.; Beverly, MA) is ligated in, yielding pCGN1125 (LB-35S-RB).

Plasmid pCGN1125 is digested with *Hind*III and *Bgl*II to 15 delete the 0.72 kb fragment of the 35S promoter. This digest is ligated with *Hind*III- and *Bam*HI-digested Cam<sup>r</sup> vector, pCGN786 (described in co-pending U.S. application Serial No. 07/382,803, filed July 19, 1989). The resulting 3.22 kb plasmid, pCGN1128, contains the 35S promoter with a 20 3' multilinker in a Cam<sup>r</sup> backbone.

Plasmid pCGN1128 is digested with *Hind*III, treated with Klenow fragment to blunt the ends and ligated with *Bgl*II linkers to yield pCGN1129, thus changing the *Hind*III site located 5' to the 35S promoter into a *Bgl*II site.

25       B.     Transit Peptide Joined to GlaA Gene

Plasmid pCGN1115 is digested with *Sal*II to remove a 1.6 kb fragment containing the SSU leader plus 48 bp of the mature SSU gene and the aroA gene. An *Xba*I digest of pCGN1129 opens the plasmid 3' to the 35S promoter.

Ligation of these two digests yields the 4.8 kb plasmid pCGN1132, containing 5'-35S-SSU leader plus 48 bp of mature SSU-*aroA*.

Plasmid pGlgA-2 is digested with *Bgl*III and *Sal*I and 5 ligated to pCGN1132 that has been digested with *Bam*HI and *Sal*I. A clone containing 5' 35S-SSU+48bp-glgA 3' is selected and designated pCGN1439.

C. Transit Peptide Joined to GlgC Gene

Plasmid pGlgC-37 is digested with *Bgl*III and *Sal*I and 10 ligated to pCGN1132 that has been digested with *Bam*HI and *Sal*I. A clone containing 5' 35S-SSU+48bp-glgC 3' is selected and designated pCGN1440.

15 Example 3 Cloning of Patatin Regulatory Regions  
and Preparation of Patatin-5'-nos-3'  
Expression Cassettes

This example describes the cloning of a patatin-5' regulatory region from potato and the preparation of 20 patatin-5'-nos-3' expression cassette pCGN2143.

Genomic DNA is isolated from leaves of *Solanum tuberosum* var. Kennebec as described in Dellaporta et al., *Plant Mol. Biol. Reporter* (1983) 1(4):19-21), with the following modifications: approximately 9g fresh weight of 25 leaf tissue is ground, a polytron grinding is not performed and in the final step the DNA is dissolved in 300μl of 10mM Tris, 1mM EDTA, pH 8.

A synthetic oligonucleotide, pat1, containing digestion sites for *Nhe*I, *Pst*I and *Xho*I with 24 bp of 30 homology to the 5'-region of a 701 bp fragment (coordinates 1611 to 2313) 5' to a class I patatin gene, isolated from

*Solanum tuberosum* var. Maris Piper (Bevan et al., NAR (1986) 14:4625-4638), is synthesized (Applied BioSystems 380A DNA synthesizer). A second synthetic oligonucleotide, pat2, containing digestion sites for *Bam*HI and *Spe*I with 25 5 bp of homology to the 3' region of the 703 bp piece is also synthesized.

Using the genomic potato DNA as a template, and pat1 and pat2 as primers, a polymerase chain reaction (PCR) is performed in a Perkin-Elmer/Cetus thermal cycler with the 10 manufacturer's reagents and in accordance with the manufacturer's instructions. The reaction contains 62.5 $\mu$ l H<sub>2</sub>O, 10 $\mu$ l 10X Reaction buffer, 16 $\mu$ l dNTP's [1.25mM dCTP, dATP, dGTP & dTTP], 5 $\mu$ l pat1 (20mM), 5 $\mu$ l pat2 (20mM), 1 $\mu$ l potato genomic DNA (3 $\mu$ g/ $\mu$ l), 0.5 $\mu$ l Tag polymerase. The PCR 15 is performed for 25 cycles with melting for 1 minute at 94°C, annealing for 2 minutes at 37°C and chain elongation for 3 minutes at 72°C. The resulting PCR product fragments (approximately 700 bp) are digested with *Nhe*I and *Bam*HI. Plasmid pCGN1586N ('5-D35S-TMVΩ'-nos-3'; pCGN1586 20 (described below) having a *Nhe*I site 5' to the 35S region) is digested with *Nhe*I and *Bam*HI to delete the D35S-Ω' fragment. Ligation of *Nhe*I-*Bam*HI digested pCGN1586N, which contains the nos-3' region, and the PCR fragments yields a 25 patatin-5'-nos3' cassette with *Spe*I, *Bam*HI, *Sal*I and *Sst*I restriction sites between the 5' and 3' regions for insertion of a DNA sequence of interest.

The 5' region of a clone, designated pCGN2143 is sequenced. Plasmid pCGN2143 has a Kennebec patatin-5'

region that is 702 bp in length and 99.7% homologous to the native sequence (as reported by Bevan (1986) *supra*).

Synthetic oligonucleotides, pat5 and pat6, are prepared as described above. Pat5 and pat6 contain complementary sequences which contain the restriction digest sites *Nhe*I, *Xho*I and *Pst*I. Pat5 and pat6 are annealed to create a synthetic linker. The annealed linker is ligated to pCGN2143 that has been linearized with *Eco*RI and treated with Klenow polymerase to generate blunt ends.

10 A plasmid, pCGN2162 which has the following restriction sites at the 3' end of nos is selected:

5'-*Eco*RI-*Nhe*I-*Xho*I-*Pst*I-*Eco*RI.

Construction of pCGN1586/1586N

Plasmid pCGN2113 (6.1 kb) contains a double-35S promoter (D35S) and the *tml*-3' region with multiple cloning sites between them, contained in a pUC-derived plasmid backbone bearing an ampicillin resistance gene (*Amp*<sup>R</sup>). The promoter/ *tml* cassette is bordered by multiple restriction sites for easy removal. Plasmid pCGN2113 is digested with 20 *Eco*RI and *Sac*I, deleting the 2.2 kb *tml*-3' region. Plasmid pBI221.1 (Jefferson, R.A., *Plant Mol. Biol. Reporter* (1987) 5:387-405) is digested with *Eco*RI and *Sac*I to delete the 0.3 kb nos-3' region. The digested pCGN2113 and pBI221.1 DNAs are ligated together, and the resultant 4.2 kb 25 recombinant plasmid with the *tml*-3' of pCGN2113 replaced by nos-3' is designated pCGN1575 (5'-D35S-nos-3').

Plasmid pCGN1575 is digested with *Sph*I and *Xba*I, blunt ends generated by treatment with Klenow fragment, and the ends are ligated together. In the resulting plasmid,

pCGN1577, the *Sph*I, *Pst*I, *Sal*I and *Xba*I sites 5' of the D3S5 promoter are eliminated.

Plasmid pCGN1577 is digested with EcoRI, the sticky ends blunted by treatment with Klenow fragment, and 5 synthetic BglII linkers (d(pCAGATCTG) New England Biolabs, Inc.; Beverly, MA) are ligated in. A total of three BglII linkers are ligated into the EcoRI site creating two PstI sites. The resulting plasmid, termed pCGN1579 (D35S-nos-3'), has a 3' polylinker consisting of 5'-EcoRI, BglII, 10 PstI, BglII, PstI, BglII, EcoRI-3'.

A tobacco mosaic virus omega' (TMV $\Omega'$ ) region (Gallie et al., NAR (1987) 15(21):8693-8711) with *Bgl*III, *Nco*I, *Bam*HI, *Sal*I and *Sac*I restriction sites:

15 BglIII  
 5'-CAGGAGATCT TATTTTACA ACAATTACCA ACAACAAACAA ACAACAAACA  
 ACATTACAAT TACTATTTAC AATTACACCA TGGATCCGTC GACGAGCTC 3'  
 NcoI BamHI SalI SacI  
 20 (SEQ ID NO: 5)

is synthesized on an Applied Biosystems® 380A DNA synthesizer and digested with *Bgl*II and *Sac*I. Plasmid pCGN1577 is digested with *Bam*HI and *Sac*I and the synthetic TMVΩ' is ligated in between the 5'-D35S and nos-3' regions. The resulting plasmid is designated pCGN1586 (5'-D35S-TMVΩ'-nos-3'). Plasmid pCGN1586N is made by digesting pCGN1586 with *Hind*III and filling in the 5' overhang with Klenow fragment, thus forming a *Nhe*I site 5' to the D35S region.

pCGN2143 is also described in co-pending U.S. application Serial No. 07/536,392 filed June 11, 1990, which is hereby incorporated by reference.

#### 5 Example 4 Preparation of Binary Vectors

This example describes the construction of a binary vector containing: (1) the patatin-5' region from *Solanum tuberosum* var. Kennebec, (2) DNA encoding a transit peptide from soybean RuBisCo SSU protein, (3) 48 bp of DNA encoding 10 16 amino acids of mature RuBisCo SSU protein from pea, (4) the *glgA* coding region from *E. coli* 618 and (5) the nos-3' region.

##### A. Glga Construct

Plasmid pCGN2162 prepared as described in Example 3 is 15 digested with *SpeI* and *SalI*, opening the plasmid between the patatin-5' region and nos-3' region. Plasmid pCGN1439 (described in Example 2) is digested with *XbaI* and *SalI* and ligated with pCGN2162 to yield pCGN1454. Plasmid pCGN1454 consists of 5'-Kennebec patatin-SSU+48-glgA-nos3'.

20 pCGN1454 is digested with *XhoI* and treated with Klenow polymerase to generate blunt ends. pCGN1557 is digested with *XbaI* and treated with Klenow polymerase to generate blunt ends. The fragments resulting from the digests are ligated together. The transformation is plated onto ECLB 25 containing gentamycin, IPTG and X-Gal. White colonies are picked and screened for ampicillin sensitivity. Gent<sup>r</sup>, Amp<sup>s</sup> clones are analyzed and two clones are selected. pCGN1457 has the 5'patatin-SSU+48bp-glgA-nos3' inserted into pCGN1557 such that it transcribes in the opposite

direction from the 35S-Kan<sup>r</sup>-tml gene. pCGN1457B has the 5'patatin-SSU+48bp-glgA-nos3' inserted into pCGN1557 such that it transcribes in the same direction as the 35S-Kan<sup>r</sup>-tml gene.

5           B.     GlgC Construct

Plasmid pCGN2162 prepared as described in Example 3 is digested with *Spe*I and *Sal*I, opening the plasmid between the patatin-5' region and nos-3' region. Plasmid pCGN1440 (described in Example 2) is digested with *Xba*I and *Sal*I and 10 ligated with pCGN2162 to yield pCGN1453. Plasmid pCGN1453 consists of 5'-Kennebec patatin-SSU+48-glgC-nos3'.

pCGN1453 is digested with *Pst*I and ligated to a *Pst*I digest of pCGN1557. The transformation is plated as described above and colonies are screened for ampicillin 15 sensitivity. Gent<sup>r</sup>, Amp<sup>s</sup> clones are analyzed and one clone, pCGN1455, is selected. pCGN1455 has the 5'patatin-SSU+48bp-glgC-nos3' inserted into pCGN1557 such that it transcribes in the same direction as the 35S-Kan<sup>r</sup>-tml gene.

C.     Construction of pCGN1557

20           Plasmid pCGN1557 (McBride and Summerfelt, *Plant Mol. Biol.* (1990) 14(27):269-276) is a binary plant transformation vector containing the left and right T-DNA borders of *Agrobacterium tumefaciens* octopine Ti-plasmid pTiA6 (Currier and Nester, *J. Bact.* (1976) 126:157-165), 25 the gentamicin resistance gene (Gen<sup>r</sup>) of pPH1JI (Hirsch and Beringer, *Plasmid* (1984) 12:139-141), an *Agrobacterium rhizogenes* Ri plasmid origin of replication from pLJbB11 (Jouanin et al., *Mol. Gen. Genet.* (1985) 201:370-374), a 35S promoter-Kan<sup>r</sup>-tml-3' region capable of conferring

kanamycin resistance to transformed plants, a ColE1 origin of replication from pBR322 (Bolivar et al. (1977) *supra*) and a lacZ' screenable marker gene from pUC18 (Yanisch-Perron et al., (1985) *supra*). The construction of pCGN1557 5 is also described in co-pending U.S. application Serial No. 07/494,722, filed March 16, 1990.

#### **Example 5 Preparation of Transgenic Plants**

This example describes the transformation of 10 *Agrobacterium tumefaciens* with glycogen biosynthetic enzyme gene nucleic acid constructs in accordance with the present invention and the cocultivation of these *A. tumefaciens* with plant cells to produce transgenic plants containing the glycogen constructs.

15       A. Transformation of Agrobacterium tumefaciens

Cells of *Agrobacterium tumefaciens* strain 2760 (also known as LBA4404, Hoekema et al., *Nature* (1983) 303:179-180) are transformed with binary vectors, such as pCGN1457, pCGN1457B and pCGN1455 (as described in Example 4) using 20 the method of Holsters, et al., (*Mol. Gen. Genet.*, (1978) 163:181-187). The transformed *A. tumefaciens* are then used in the co-cultivation of plants.

The *Agrobacterium* are grown on AB medium (K<sub>2</sub>HPO<sub>4</sub> 6g/l, NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O 2.3 g/l, NH<sub>4</sub>Cl 2g/l, KC1 3g/l, glucose 5g/l, 25 FeSO<sub>4</sub> 2.5mg/l, MgSO<sub>4</sub> 246mg/l, CaCl<sub>2</sub> 14.7mg/l, 15g/l agar), plus 100µg/l gentamycin sulfate and 100µg/l streptomycin sulfate for 4-5 days. Single colonies are inoculated into 10ml of MG/L broth (per liter: 5g mannitol, 1g L-Glutamic acid or 1.15g sodium glutamate, 0.5g KH<sub>2</sub>PO<sub>4</sub>, 0.10g NaCl,

0.10g MgSO<sub>4</sub>·7H<sub>2</sub>O, 1μg biotin, 5g tryptone, 2.5g yeast extract; adjust pH to 7.0) and are incubated overnight in a shaker at 30°C and 180rpm. Prior to co-cultivation, the *Agrobacterium* culture is centrifuged at 12,000xg for 10 minutes and resuspended in 20ml of MS medium (#510-1118, Gibco; Grand Island, NY).

B. Cocultivation with Potato Cells

Feeder plates are prepared by pipetting 0.5ml of a tobacco suspension culture (~10<sup>6</sup>cells/ml) onto 0.8% agar 10 co-cultivation medium, containing Murashige and Skoog salts (#510-117, Gibco; Grand Island, NY), thiamine-HCl (1.0 mg/l), nicotinic acid (0.5mg/l), pyridoxine HCl (0.5mg/l), sucrose (30g/l), zeatin riboside (5μM), 3-indoleacetyl-DL-aspartic acid (3μM), pH 5.9. The feeder plates are 15 prepared one day in advance and incubated at 25°C. A sterile 3mm filter paper disk is placed on top of the tobacco cells after the suspension cells have grown for one day.

Tubers of *Solanum tuberosum* var Russet Burbank between 20 the age of 1 and 6 months post harvest are peeled and washed in distilled water. All subsequent steps are carried out in a flow hood using sterile techniques. For surface sterilization, tubers are immersed in a solution of 10% commercial bleach (sodium hypochlorite) with 2 drops of 25 Ivory® liquid soap per 100ml for 10 minutes. Tubers are rinsed six times in sterile distilled water and kept immersed in sterile liquid MS medium (#1118, Gibco; Grand Island; New York) to prevent browning.

Tuber discs (1-2mm thick) are prepared by cutting columns of potato tuber with a ~1cm in diameter cork borer and slicing the columns into discs of the desired thickness. Discs are placed into the liquid MS medium 5 culture of the transformed *Agrobacterium tumefaciens* containing the binary vector of interest ( $1 \times 10^7$ - $1 \times 10^8$  bacteria/ml) until thoroughly wetted. Excess bacteria are removed by blotting discs on sterile paper towels. The discs are co-cultivated with the bacteria for 48 hours on 10 the feeder plates and then transferred to regeneration medium (co-cultivation medium plus 500mg/l carbenicillin and 100mg/l kanamycin). In 3 to 4 weeks, shoots develop from the discs.

When shoots are approximately 1 cm, they are excised 15 and transferred to a 0.8% agar rooting medium containing MS salts, thiamine-HCl (1.0mg/l), nicotinic acid (0.5mg/l), pyridoxine-HCl (0.5mg/l), sucrose (30g/l), carbenicillin (200mg/l) and kanamycin (100-200 mg/l) pH 5.9. Plants are rooted two times with at least one rooting taking place on 20 rooting medium with the higher level of kanamycin (200mg/l). Plants which have rooted twice are then confirmed as transformed by performing NPTII blot activity assays (Radke, S.E. et al, *Theor, Appl. Genet.* (1988) 75:685-694). Plants which are not positive for NPTII 25 activity are discarded.

**Example 6      Analysis of Tubers from Transformed Potato Plants**

30      In this Example, measurement of specific gravity in tubers from transgenic potato plants is described.

Rooted plants, transformed as described in Example 5, are cut into five sections at the internodes and each section is rooted again, also as described in Example 5. The newly rooted plants are transplanted from rooting medium to soil and placed in a growth chamber (21°C, 16 hour days with 250-300  $\mu\text{E}/\text{m}^2/\text{sec}$ ). Soil is prepared as follows: For about 340 gallons, combine 800 pounds 20/30 sand (approximately 14 cubic feet), 16 cubic feet Fisons Canadian Peat Moss, 16 cubic feet #3 vermiculite, and approximately 4.5 pounds hydrated lime in a Gleason mixer. The soil is steamed in the mixer for two hours; the mixer mixes for about 15 seconds at intervals of fifteen minutes over a period of one hour to ensure even heating throughout the soil. During and after the process of steaming, the soil reaches temperatures of at least 180°F for one hour. The soil is left in the mixer until the next day. At that time, hydrated lime is added, if necessary, to adjust the pH to range between 6.30 and 6.80.

The relative humidity of the growth chamber is maintained at 70-90% for 2-4 days, after which the humidity is maintained at 40-60%. When plants are well established in the soil, after approximately two weeks, they are transferred to a greenhouse. In the greenhouse, plants are grown in 6.5 inch pots in a soil mix of peat:perlite:vermiculite (11:1:9), at an average temperature of 24°C day/12°C night. Day length is approximately 12 hours and light intensity levels vary from approximately 600 to 1000  $\mu\text{E}/\text{m}^2/\text{sec}$ .

Tubers from each plant are harvested and washed 14 weeks after transfer to the greenhouse. Immediately after harvest, three to five uniformly sized tubers from each pot are weighed and their specific gravity determined. In 5 determining specific gravity, the tubers from each plant are first collectively weighed in air and then collectively weighed in water. Specific gravity is determined, where  $x$  = the weight of tubers in air and  $y$  = the weight of tubers in water, as  $x/(x - y)$ .

10 In general, the specific gravities of tubers from five replicates of plants transformed with the *glgA* constructs (pCGN1457 and pCGN1457B) and of tubers from control plants are determined. Control plants include regenerated non-transformed potato plants and transgenic potato plants 15 which lack the *glgA* constructs. Controls are subjected to the transformation and regeneration culture and growth conditions described above in production of *glgA* transformed plants. To compare values from each tuber sample, the specific gravity measurements are converted to 20 reflect % total solids content of tubers. Percent total solids is calculated as (specific gravity)  $\times$  (199.63) - 194.84 (Porter, et al., *Am. Pot. J.* (1964) 41:329-336). Differences are detected in percent total solids as 25 determined for tubers from several of the *glgA* transformed plants as compared to tubers from control plants.

Results are presented in Table 1 which represent average specific gravity of tubers of 5 replicate plants, except as otherwise indicated. Specific gravity measurements were determined for three to five uniformly

sized tubers from each plant and the measurements of the tubers from the replicate plants were then averaged to determine average specific gravity (SpGr) of tubers for each transformation event. Values for one set of 5 transformed control plants (Tx) and one set of untransformed/regenerated control plants (Rg) for each construct are shown at the top of their respective columns. Transformed control plants were transformed with a non-carbohydrate-related gene.

10

Table 1  
Average Specific Gravity Measurements

	<u>Event</u>	<u>SpGr</u>	<u>Event</u>	<u>SpGr</u>
15	Controls		Controls	
	Tx	1.079	Tx	1.083
	Rg	1.081	*Rg	1.077
	Transformed Plants		Transformed Plants	
	1457- 3	1.073	1457B- 3	1.062
20	1457- 4	1.060	1457B- 4	1.075
	1457- 6	1.076	1457B- 5	1.073
	1457- 7	1.080	1457B- 7	1.066
	1457- 8	1.077	1457B- 8	1.066
	1457- 9	1.067	1457B- 9	1.063
25	1457-10	1.083	1457B-10	1.075
	1457-11	1.065	1457B-12	1.065
	1457-12	1.066	1457B-13	1.058
	1457-13	1.080	*1457B-15	1.053
	1457-14	1.062	1457B-16	1.075
30	1457-15	1.064	1457B-17	1.053
	1457-16	1.068	1457B-18	1.068
	1457-17	1.069	1457B-21	1.081
	1457-18	1.060	1457B-22	1.067
	1457-19	1.069	1457B-23	1.069
35	1457-20	1.066	1457B-24	1.068
	1457-22	1.068		

\*Only 4 replicate plants were available for these samples.

It is readily apparent from the data presented in Table 1 that transgenic plants were obtained which produce tubers having an altered specific gravity as compared to the tubers from control plants.

5        Statistical analysis was conducted on the specific gravity measurements of tubers from the 5 replicates of one of the transformation events as compared to the specific gravity measurements of tubers from two control events. The event analyzed was 1457-4 which had an average specific  
10      gravity of 1.060. The specific gravity measurements of tubers from the individual replicates that were used to calculate the average for this event were 1.059, 1.057, 1.067, 1.066, and 1.053. The specific gravity measurements for replicates of control tubers were as follows. Tx (ave.  
15      1.079): 1.076, 1.082, 1.073, 1.083, and 1.079. Rx (ave.  
1.081): 1.076, 1.087, 1.083, 1.082, and 1.077. These measurements were converted to percent solids as described above and the percent solids values were used for statistical analysis as follows.

20       A comparison of sample means was conducted on the percent solids values calculated for the three events, 1457-4, Tx and Rx, by calculating the t value (Student's t) and determining statistical difference based on a standard table of values for t. (See, for example, Steel and Torrie  
25      (1980) *Principles and Procedures of Statistics: A Biometrical Approach* (McGraw-Hill pub.) Chapter 5 and Table A.3). These analyses indicate a significant difference between the average specific gravity measurements of transgenic tubers as compared to control tubers at a

confidence level of greater than 99%. The average specific gravity measurements of the two control groups were not significantly different.

Further analysis may be conducted on tubers from 5 selected PCGN1457 and PCGN1457B transformed plants and from non-transformed controls (RB-43) to determine starch content, amylose percentages and to elucidate chain length distribution in the amylopectin component of the starch. Starch granules are isolated as described by Boyer et al. 10 (1976) *Cereal Chemistry* 53:327-337) and starch content estimated on a weight basis (starch wt/ fresh wt). Amylose percentages are determined by gel-filtration analysis (Boyer et al. (1985) *Starch/Starke* 37:73-79). Chain length distribution patterns are determined by HPLC analysis as 15 described by Sanders et al. (1990) *Cereal Chemistry* 67:594-602). Amylopectins are characterized by the ratios (on a weight basis) of low molecular weight chains to high molecular weight chains as described by Hizukuri (*Carbohydrate Research* (1985) 141:295-306). Results of 20 these analyses are presented in Table 2.

Table 2  
Analyses of Transgenic Potato Tuber Starch

		<u>Construct</u>	<u>Spec. Gravity</u>	<u>Starch</u>	<u>Amylose</u>	<u>Chains</u>	% High	% Low	% Low	M.W./ High
5	10						M.W.	M.W.	M.W.	M.W.
10	RB-43 1457-4 1457-17 1457-18	RB-43	1.081	17.1	23	33	66	20	80	4.0
		1457-4	1.060	11.0	12	20	72	28	85	2.6
		1457-17	1.069	14.6	24	15	85	5.7	84	5.7
		1457-18	1.060	11.8	8	26				
15	RB-43 1457B-15 1457B-17	RB-43	1.077	17.2	27	15	85	5.7	84	3.2
		1457B-15	1.053	9.0	9	26				
		1457B-17	1.053	12.5	19					

The data presented in Table 2 indicate that tubers from transgenic plants which have an altered specific gravity, also have altered starch. In particular, the percentage of amylose in the transgenic potato tubers is decreased. In addition, the amylopectin portion of the starch from transgenic potato tubers has more low molecular weight chains and less high molecular weight chains than wild type potato tuber amylopectin, thus indicating that the amylopectin from transgenic tubers has more branch points.

It is evident from the above results, that plant cells and plants can be produced which have improved properties or may produce a desired product. In accordance with the subject invention, it is now seen that glycogen biosynthesis enzyme sequences may be introduced into a plant host cell and be used to express such enzyme or enzymes or to modify native starch precursors. Moreover, it is seen that such enzymes demonstrate biological activity on plant starch precursors resulting in a

demonstrable phenotype *in planta*, namely altered specific gravity. In addition, the activity of glycogen biosynthetic enzymes in plants has been shown to result in starch having altered properties, in particular altered 5 ratios of amylose/amylopectin and altered distribution of low molecular weight chain lengths to high molecular weight chain lengths in the amylopectin fraction. In this manner, plants, including plant cells and plant parts, having modified starch properties may be obtained, wherein the 10 modified starch has unique and desirous properties.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were 15 specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be readily apparent 20 to those of ordinary skill in the art in light of the teaching of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

What is claimed is:

1. A nucleic acid construct comprising a glycogen biosynthesis enzyme sequence and an endogenous plant sequence.
- 5 2. The construct of Claim 1 wherein said endogenous plant sequence comprises a plastid transit peptide.
3. The construct of Claim 1 wherein said enzyme sequence comprises in the 5'-3' direction a sense sequence.
4. The construct of Claim 1 wherein said enzyme sequence comprises a glycogen synthase sequence.
- 10 5. The construct of Claim 1 wherein said enzyme sequence comprises ADP-glucose pyrophosphorylase sequence.
6. The construct of any one of Claims 1-5 wherein said sequences are DNA.
- 15 7. The construct of Claim 4 or 5 having a DNA sequence in the 5'-3' direction of transcription, a transcription initiation region, a translational initiation region, an open reading frame sequence encoding said enzyme, and a transcription/ translational termination region, wherein said regulatory regions are capable of functioning in a plant cell.
- 20 8. The construct of Claim 7 further comprising a plastid transit peptide immediately upstream of said open reading frame.
- 25 9. The construct of Claim 7 wherein said transcription initiation region comprises at least a portion of the regulatory region 5' to a patatin gene from *Solanum tuberosum*.

10. The construct of Claim 7 further comprising a T-DNA border.

11. A glycogen biosynthesis enzyme joined to a plastid transit peptide.

5 12. A plant cell comprising a glycogen biosynthesis enzyme.

13. The plant cell of Claim 12 comprising glycogen synthase.

14. The plant cell of Claim 12 wherein said enzyme  
10 is present in an amyloplast.

15. The plant cell of Claim 12 selected from the group consisting of corn, sorghum, rice, potato, tapioca, cassava, arrowroot, sago and cereal grain crops.

16. A plant comprising a cell of any one of Claims  
15 12-14.

17. A method to modify a starch storage organ comprising

growing a plant, under conditions that will permit the formation of a starch storage organ, having in  
20 its genome an expression construct which is capable of expressing a glycogen biosynthesis enzyme preferentially in a starch storage organ.

18. The method of Claim 17 wherein said starch storage organ is a tuber having a decreased specific  
25 gravity.

19. A plant starch storage organ having a modified specific gravity produced by the method of Claim 18.

20. The method of Claim 17 wherein said starch storage organ comprises modified starch.

21. The method of Claim 20 wherein said modified starch has an altered amylose to amylopectin ratio as compared to that of a control starch storage organ.

22. The method of Claim 20 wherein said modified starch has an altered ratio of low to high molecular weight chains in the amylopectin fraction as compared to that of a control start storage organ.

23. A plant starch storage organ having modified starch produced according to any one of Claims 20-22.

GLGA-40

GATCTAACAG GAGCGATAAT GCAGGGTTTA CATGTATGTT CAGAGATGTT CCCGGCTGCTT 60  
AAAACCCGGG GTCTGGCTGA TGTATTGGG GCATTACCCG CAGCACAAAT CGCAGACGGC 120  
GTTGACCGCTC CGGTACTGTT GCCTGCATT CCCGATATT GCCTGGCGT GACCGATGCC 180  
CAGGTAAT CCCGTCGTGA TACCTTCGCC GGACATATCA CGCTGTTGTT CGGTCAATTAC 240  
AACGGGGTTG GCATTTACCT GATTGACGGG CCGCATCTCT ATGATCGTCC GGGAAAGCCCCG 300  
TATCAGGATA CCAACTTATT TGCCTATAACC GACAACGTAT TGGCTTTGTC GCTGCTGGGG 360  
TGGGTGGGG CAGAAATGGC CAGCGGGCTT GACCCATTCT GGCGTCCTGA TGTGGTGCAT 420  
GGGCACCACT GGCATGCAGG CCTTGCGCCT GCGTATCTGC CGGCCGGGG GCGTCCGGGG 480  
AAGTCGGTGT TTACTGGCA AACCTGGCC TATCAAAGGCA TGTTTATGC ACATCACATG 540

FIG. 1A

7:  
3

AATGACATCC AATTGCCATG GTCATTCTT AATATTCATG GGCTGGAATT CAACGGACAA 600  
ATCTCTTTCC TGAAGGCCGG TCTGTACTAT GCCGATCACA TTACGGCGGT CAGTCCAACC 660  
TACGGCTCGCG AGATCACCGGA ACCGCAGTTT GCCTACGGTA TGGAAAGGTCT GTTGCAACAG 720  
CGTCACCGTG AAGGGCGTCT TTCCGGCGTA CTGAAACGGCG TGGACCGAGAA AATCTGGAGT' 780  
CCAGAGACGG ACTTACTGTT GCCCTCGCGT TACACCCGG ATACGTTGGA AGATAAACCG 840  
GAAATAAAGC GCCAGTTACA ATCGCAATG GGGCTTAAGG TTGACCGATAAA AGTGCCGCTT 900  
TTTGCAAGTGG TGAGCCGTCT GACCAGGCCAG AAAGGTCTCG ACCTGGTGT GGAAAGCCCTA 960  
CCGGGGCTTC TGGAGCAGGG CGGGCAGGCTG GCGCTACTCG GCGGGGGGA TCCGGTGCTG 1020  
CAGGAAAGGTT TCCTTGCGGC GGCAGGGAA TACCCCCGGTC AGGTGGCCGT TCAGATTGGCC 1080

FIG. 1B

3  
13

TATCACCGAAG CATTTCGCA TCGCATTATG GGCGGGCGG ACGTCAATTCT GGTGCCAGC 1140  
CGTTTGAAAC CGTGGGGCTT AACGCAAACCTT TAGGATTGA AGTACGGTAC GCTGCCGTTA 1200  
GTGCGGGCA CGGGTGGCT TGCTGATACT GTTCTGACT GTTCTCTTGA GAACCTTGCA 1260  
GATGGCGTCG CCAGTGGGTT TGTCTTGTGAA GATAGTAATG CCTGGTCGCT GTTACGGCT 1320  
ATTCGACGTG CTTTGTACT GTGGTCCCGT CCTCACTGT GGCGGTTGT GCAACGTCAG 1380  
GCTATGGCAA TGGATTTAG CTGGCAGGTC GCGGCGAAGT CGTACCGTGA GCTTTACTAT 1440  
CGCTCGAAAT AGTTTCACT CGAC 1464

FIG. 1C

**GLGA-40**

MET Gln Val Leu His Val Cys Ser Glu MET Phe Pro Leu Leu Lys Thr  
 1 5 10 15

Gly Gly Leu Ala Asp Val Ile Gly Ala Leu Pro Ala Ala Gln Ile Ala  
 20 25 30

Asp Gly Val Asp Ala Arg Val Leu Leu Pro Ala Phe Pro Asp Ile Arg  
 35 40 45

Arg Gly Val Thr Asp Ala Gln Val Val Ser Arg Arg Asp Thr Phe Ala  
 50 55 60

Gly His Ile Thr Leu Leu Phe Gly His Tyr Asn Gly Val Gly Ile Tyr  
 65 70 75 80

Leu Ile Asp Ala Pro His Leu Tyr Asp Arg Pro Gly Ser Pro Tyr His  
 85 90 95

Asp Thr Asn Leu Phe Ala Tyr Thr Asp Asn Val Leu Arg Phe Ala Leu  
 100 105 110

**FIG. 2A**

Leu Gly Trp Val Gly Ala Glu MET Ala Ser Gly Leu Asp Pro Phe Trp  
115 120 125

Arg Pro Asp Val Val His Ala His Asp Trp His Ala Gly Leu Ala Pro  
130 135 140

Ala Tyr Leu Ala Ala Arg Gly Arg Pro Ala Lys Ser Val Phe Thr Gly  
145 150 155 160

His Asn Leu Ala Tyr Gln Gly MET Phe Tyr Ala His His MET Asn Asp  
165 170 175

Ile Gln Leu Pro Trp Ser Phe Phe Asn Ile His Gly Leu Glu Phe Asn  
180 185 190

Gly Gln Ile Ser Phe Leu Lys Ala Gly Leu Tyr Tyr Ala Asp His Ile  
195 200 205

Thr Ala Val Ser Pro Thr Tyr Ala Arg Glu Ile Thr Glu Pro Gln Phe  
210 215 220

FIG. 2B

Ala Tyr Gly MET Glu Gly Leu Gln Gln Arg His Arg Glu Gly Arg  
225 230 235 240

Leu Ser Gly Val Leu Asn Gly Val Asp Glu Lys Ile Trp Ser Pro Glu  
245 250 255

Thr Asp Leu Leu Ala Ser Arg Tyr Thr Arg Asp Thr Leu Glu Asp  
260 265 270

Lys Ala Glu Asn Lys Arg Gln Leu Gln Ile Ala MET Gly Leu Lys Val  
275 280 285

Asp Asp Lys Val Pro Leu Phe Ala Val Val Ser Arg Leu Thr Ser Gln  
290 295 300

Lys Gly Leu Asp Leu Val Leu Glu Ala Leu Pro Gly Leu Leu Glu Gln  
305 310 315 320

Gly Gly Gln Leu Ala Leu Leu Gly Ala Gly Asp Pro Val Leu Gln Glu  
325 330 335

FIG. 2C

Gly Phe Leu Ala Ala Ala Ala Glu Tyr Pro Gly Gln Val Gly Val Gln  
340 345 350

Ile Gly Tyr His Glu Ala Phe Ser His Arg Ile MET Gly Gly Ala Asp  
355 360 365

Val Ile Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Thr Gln Leu  
370 375 380

Tyr Gly Leu Lys Tyr Gly Thr Leu Pro Leu Val Arg Arg Thr Gly Gly  
385 390 395 400

Leu Ala Asp Thr Val Ser Asp Cys Ser Leu Glu Asn Leu Ala Asp Gly  
405 410 415

Val Ala Ser Gly Phe Val Phe Glu Asp Ser Asn Ala Trp Ser Leu Leu  
420 425 430

Arg Ala Ile Arg Arg Ala Phe Val Leu Trp Ser Arg Pro Ser Leu Trp  
435 440 445

FIG. 2D

$\delta_i \beta$

Arg Phe Val Gln Arg Gln Ala MET Ala MET Asp Phe Ser Trp Gln Val  
450 455 460

Ala Ala Lys Ser Tyr Arg Glu Leu Tyr Tyr Arg Ser Lys  
465 470 475

FIG. 2E

GATCTAGGAG CGATA ATG GTT AGT TTA GAG AAG AAC GAT CAC TCA ATG  
 MET Val Ser Leu Glu Lys Asn Asp His Leu MET  
 1 5 10

48

TTG GCG CGC CAG CTG CCA TTG AAA TCT GTT GCC CTC ATA CTG GCG GGA  
 Leu Ala Arg Gln Leu Pro Leu Lys Ser Val Ala Leu Ile Leu Ala Gly  
 15 20 25

96

GGA CGT GGT ACC CGC CTG AAG GAT TTA ACC AAT AAG CGA GCA AAA CCG  
 Gly Arg Gly Thr Arg Leu Lys Asp Leu Thr Asn Lys Arg Ala Lys Pro  
 30 35 40

144

GCC GTA CAC TTC GGC GGT AAG TTC CGC ATT ATC GAC TTT GCC CTG TCT  
 Ala Val His Phe Gly Gly Lys Phe Arg Ile Arg MET Gly Val Ile Thr Gln Tyr  
 45 50 55

192

AAC TGC ATC MAC TCC GGG ATC CGT ATG GGC GTG ATC ACC CAG TAC  
 Asn Cys Ile Asn Ser Gly Ile Arg Arg MET Gly Val Ile Thr Gln Tyr  
 60 65 70 75

240

CAG TCC CAC ACT CTG GTG CAG CAC ATT CAG CGC GGC TGG TCA TTC TAC  
 Gln Ser His Thr Leu Val Gln His Ile Gln Arg Gly Trp Ser Phe Phe  
 80 85 90

288

FIG. 3A

16  
15

ATT GAA GAA ATG AAC GAG TTT GTC GAT CTG CTG CCA GCA CAG CAG AGA	336
Asn Glu Glu MET Asn Glu Phe Val Asp Leu Pro Ala Gln Gln Arg	105
ATG AAA GGG GAA AAC TGG TAT CGC GGC ACC GCA GAT GCG GTC ACC CAA	384
MET Lys Glu Lys Trp Tyr Arg Gly Thr Ala Asp Ala Val Thr Gln	110 115 120
AAC CTC GAC ATT ATC CGC CGT TAT AAA GCG GAA TAC GTG ATC CTG	432
Asn Leu Asp Ile Ile Arg Arg Tyr Lys Ala Glu Tyr Val Val Ile Leu	125 130 135
GCG GGC GAC CAT ATC TAC AAG CAA GAC TAC TCG CGT ATG CTT ATC GAT	480
Ala Gly Asp His Ile Tyr Lys Gln Asp Tyr Ser Arg MET Leu Ile ASP	140 145 150 155
CAC GTC GAA AAA GGC GCA CGT TGC ACC GTT GCT TGT ATG CCA GTC CCG	528
His Val Glu Lys Gly Ala Arg Cys Thr Val Ala Cys MET Pro Val Pro	160 165 170
ATT GAA GAA GCC TCC GCA TTT GGC GTT ATG GCG GTT GAT GAG AAC GAT	576
Ile Glu Glu Ala Ser Ala Phe Gly Val MET Ala Val Asp Glu Asn Asp	175 180 185

FIG. 3B

AAC	ATT	ATC	GAA	TTC	GTT	GAA	AAA	CCT	GCT	AAC	CCG	CCG	TCA	ATG	CCG	624
Lys	Ile	Ile	Glu	Phe	Val	Glu	Lys	Pro	Ala	Asn	Pro	Pro	Ser	MET	Pro	
190							195								200	
11 15																
AAC	GAT	CCG	AGC	AAA	TCT	CTG	GCG	AGT	ATG	GGT	ATC	TAC	GTC	TTT	GAC	672
Asn	Asp	Pro	Ser	Lys	Ser	Leu	Ala	Ser	MET	Gly	Ile	Tyr	Val	Phe	Asp	
205									210			215				
220																
GCC	GAC	TAT	CTG	TAT	GAA	CTG	CTG	GAA	GAA	GAT	CGC	GAT	GAG	AAC	720	
Ala	Asp	Tyr	Leu	Tyr	Glu	Leu	Leu	Glu	Glu	Asp	Asp	Arg	Asp	Glu	Asn	
220					225				230						235	
240																
TCC	AGC	CAC	GAC	TTT	GGC	AAA	GAT	TTG	ATT	CCC	AAG	ATC	ACC	GAA	GCC	768
Ser	Ser	His	Asp	Phe	Gly	Lys	Asp	Leu	Ile	Pro	Lys	Ile	Thr	Glu	Ala	
									245						250	
255																
GGT	CTG	GCC	TAT	GCG	CAC	CCG	TTC	CCG	CTC	TGC	GTA	CAA	TCC	GAC	816	
Gly	Leu	Ala	Tyr	Ala	His	Pro	Phe	Pro	Leu	Ser	Cys	Val	Gln	Ser	Asp	
									260						265	

FIG. 3C

15  
16  
17

CCG GAT	GCC GAG	CCG TAC	TGG CGC GAT	GTG GGT ACG CTG	GAA GCT TAC	864
Pro Asp	Ala Glu	Pro Tyr	Trp Arg Asp Val	Gly Thr Leu Glu Ala	Tyr	
270		275		280		
TGG AAA	GCG AAC	CTC GAT	CTG GCC TCT GTG	CCG GAA CTG GAT ATG		912
Trp Lys	Ala Asn	Leu Asp	Leu Ala Ser Val	Pro Val Pro Glu Leu Asp	MET	
285		290		295		
TAC GAT	CGC AAT	TGG CCA ATT	CGC ACC TAC AAT GAA TCA TTA CCG CCA		960	
Tyr Asp	Arg Asn	Trp Pro Ile	Arg Thr Tyr Asn Glu Ser Leu Pro Pro			
300		305		310		315
GGG AAA TTC GTG CAG GAT	CGC TCC GGT AGC CAC GGG ATG ACC CTT AAC				1008	
Ala Lys Phe Val Gln Asp	Arg Ser Gly Ser His Gly MET Thr Leu Asn					
320		325		330		
TCA CTG GTT TCC GAC GGT TGT GTG ATC TCC GGT GTG GTG CAG					1056	
Ser Leu Val Ser Asp G1y Cys Val Ile Ser Gly Ser Val Val Val Gln						
335		340		345		
TCC GTG CTC TGC TCG CGC GTG AAT TCA TGC GAC ATT GAT					1104	
Ser Val Leu Phe Ser Arg Val Arg Val Asn Ser Phe Cys Asp Ile Asp						
350		355		360		

FIG. 3D

TCC GCC GTA TTG TTA CCG GAA GTA TGG GTA GGT CGC TCG TGC CGT CTG  
 Ser Ala Val Leu Pro Glu Val Trp Val Gly Arg Ser Cys Arg Leu  
 365 370 375

CGC CGC TGC GTC ATC GAT CGT GCT TGT GTT ATT CCG GAA GGC ATG GTG  
 Arg Arg Cys Val Ile Asp Arg Ala Cys Val Ile Pro Glu Gly MET Val  
 380 385 390 395

ATT GGT GAA AAC GCA GAG GAA GAT GCA CGT CGT TTC TAT CGT TCA GAA  
 Ile Gly Glu Asn Ala Glu Glu Asp Ala Arg Arg Phe Tyr Arg Ser Glu  
 400 405 410

GAA GGC ATC GTG CTG GTA ACG CGC GAA ATG CTA CGG AAG TTA GGG CAT  
 Glu Gly Ile Val Leu Val Thr Arg Glu MET Leu Arg Lys Leu Gly His  
 415 420 425

AAA CAG GAG CGA TAATGCAGGG TCGAC  
 Lys Gln Glu Arg  
 430

1323  
 1324

FIG. 3E

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09654

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> <sup>14</sup> (if several classification symbols apply, indicate all) <sup>3</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC (S) : C12P 21/00, 21/04; C12N 9/00, 15/00; A01H 1/00 US CL : 435/69.1, 69.8, 183, 320.1; 47/58		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>4</sup>		
Classification System	Classification Symbols	
U.S.	435/69.1, 69.8, 183, 320.1; 47/58	
Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched <sup>5</sup>		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <sup>14</sup>		
Category <sup>6</sup>	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
Y	EP, A, 0,368,506 (Bridges et al.) 16 May 1990, see column 4, lines 15-31.	1-10, 17-18, 20-22
Y	Journal of Biological Chemistry, Volume 261, No. 34, issued 05 December 1986, Kumar et al., "Biosynthesis of Bacterial Glycogen," pages 16256-16259, see page 16258.	1-10, 17-18, 20-22
Y	EMBO Journal, Volume 4, No.1, issued January 1985, Schreier et al., "The use of nuclear-encolded sequences to direct the light-regulated synthesis and transport of foreign protein into plant chloroplasts," pages 25-32, see pages 26 and 28.	2, 8
Y	Plant Molecular Biology, Voluumne 9, No. 4, issued 1987, Twell et al., "The 5' flanking DNA of a patatin gene directs tuber specific expression of a chimaeric gene in potato," pages 365-375, see page 371.	9
<p>* Special categories of cited documents:<sup>16</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>2</sup>  13 MARCH 1992	Date of Mailing of this International Search Report <sup>2</sup>  08APR 1992	
International Searching Authority <sup>1</sup>  ISA/US	Signature of Authorized Officer <sup>20</sup>  David Fox, Primary Examiner	

